

157  
6/56

OIPE

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION: US/09/655,160**

DATE: 01/16/2003  
 TIME: 13:57:46

Input Set : A:\PTO.VSK.txt  
 Output Set: N:\CRF4\01162003\I655160.raw

2 <110> APPLICANT: CALIFORNIA INSTITUTE OF TECHNOLOGY  
 3        Paul, STEMPERG W.  
 4        Maureen, BARR M.  
 5 <120> TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE HOMOLOGS REQUIRED FOR MALE  
 MATING BEHAVIOR

7        IN NEMATODES AND ASSAYS BASED THEREON  
 8 <130> FILE REFERENCE: CIT1520-2  
 10 <140> CURRENT APPLICATION NUMBER: US 09/655,160  
**C--> 12 <141> CURRENT FILING DATE: 2002-12-30**  
 14 <150> PRIOR APPLICATION NUMBER: US 09/479,467  
 15 <161> PRIOR FILING DATE: 2000-01-06  
 17 <160> PRIOR APPLICATION NUMBER: US 60/115,127  
 18 <161> PRIOR FILING DATE: 1999-01-06  
 20 <170> NUMBER OF SEQ ID NOS: 16  
 21 <170> SOFTWARE: PatentIn version 3.1  
 24 <210> SEQ ID NO: 1  
 25 <211> LENGTH: 12412  
 26 <212> TYPE: DNA  
 27 <213> ORGANISM: Homo sapiens PKD-1 gene  
 28 <210> FEATURE:  
 29 <211> NAME/KEY: CDS  
 30 <212> LOCATION: (1)..(12913)  
 31 <213> OTHER INFORMATION:

**W--> 34 <400> 1**

31	atq	tgg	ccg	ggc	ccg	ccc	ctg	ggc	ctg	ggc	48	
32	Met	Pro	Pro	Ala	Ala	Pro	Ala	Arg	Leu	Ala	Leu	Gly
33												
34												
35												
36												
37												
38												
39												
40												
41												
42												
43												
44												
45												
46												
47												
48												
49												
50												
51												
52												
53												
54												
55												
56												
57												
58												
59												
60												

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION: US/09/655,160**

**DATE: 01/16/2003**  
**TIME: 13:57:46**

**Input Set : A:\PTO.VSK.txt**  
**Output Set: N:\CRF4\01162003\I655160.raw**

61	100	105	110	
63	aat tta ttt aat tta aat gaa ata aac ctg aat ggg aac ccg ttt gag			384
64	Asn Leu Phe Asn Leu Ser Glu Ile Asn Leu Ser Gly Asn Pro Phe Gln			
65	66 115 120 125			
67	tgt tac tgt ggc ctg gcg tgg ctg ccg cca tgg gcg gag gac cag			432
68	Cys Asp Cys Gly Leu Ala Trp Leu Pro Gln Trp Ala Glu Glu Gln G n			
69	6+ 130 135 140			
70	gtg ccg gtc gtc cag ccc gag gca gcc acg ttt gct ggg cct ggc tcc			480
71	Val Arg Val Val Gln Pro Glu Ala Ala Thr Cys Ala Gly Pro Gly Ser			
72	73 145 150 155 160			
74	cgt aat ccc cag cct ctg ctt gac atc ccc ttt ctg gac aat ggc ttt			528
75	Leu Ala Gly Gln Pro Leu Leu Gly Ile Pro Leu Leu Asp Ser Gly Cys			
76	77 165 170 175			
78	gtt gag gac tat gtc gcc ttc ctc ctt gag aac aac tca ggc act gtc			576
79	Gly Glu Glu Tyr Val Ala Cys Leu Pro Asp Asn Ser Ser Gly Thr Val			
80	81 180 185 190			
82	gca gca gtc ttc ttt tca gct gca ccc gca ctt ccc ccc gtc			624
83	Ala Ala Val Ser Phe Ser Ala Ala His Gln Leu Leu Gln Pro Gln			
84	85 195 200 205			
86	gcc tcc aac gca ttc tcc acc ggc ccc ggc ccc gca gca ctc			672
87	Ala Cys Ser Ala Phe Cys Phe Ser Thr Gly Gln Gln Leu Ala Ala Leu			
88	89 210 215 220			
90	tcc ggg ccc tgg tcc ctg tcc ggg ggg ggg ccc ccc tcc agt ggc			720
91	Ser Gln Gln Trp Cys Leu Cys Gly Ala Ala Gln Pro Ser Ser Ala			
92	93 225 230 235			
94	tcc ttt gcc tcc ctg tcc ctc tcc tcc ggg ccc ccc gca cct ccc gcc			768
95	Ser Phe Ala Cys Leu Ser Leu Cys Ser Gly Pro Pro Ala Pro Pro Ala			
96	97 245 250 255			
98	ccc aac ttt aac ggc ccc aac ctc ctc cag ccc gac gac ttc ccc tcc			816
99	Pro Thr Cys Arg Gly Pro Thr Leu Leu Gln His Val Phe Pro Ala Ser			
100	101 260 265 270			
102	cca ggg gcc aac ctg gtc ggg ccc ccc gca ccc ccc tcc tcc gtc cag			864
103	Pro Gly Ala Thr Leu Val Gly Pro His Gly Pro Leu Ala Ser Gly Gln			
104	105 275 280 285			
106	cta gca gcc tcc ccc aac tcc gct gcc ccc ccc ccc gtc act gac aca ccc			912
107	Leu Ala Ala Phe His Ile Ala Ala Pro Leu Pro Val Thr Asp Thr Arg			
108	109 290 295 300			
110	tgg gac ttc gga gac tgc tcc gcc ggg gtc gat gcc gct ggg ccc gtc			960
111	Trp Asp Phe Gly Asp Gly Ser Ala Gln Val Asp Ala Ala Gly Pro Ala			
112	113 305 310 315			
114	gcc tcc cat ccc tat tgg ctg ccc ccc ccc tcc ccc ggg acg gcc gtg			1008
115	Ala Ser His Arg Tyr Val Leu Pro Gly Arg Tyr His Val Thr Ala Val			
116	117 325 330 335			
118	ctg gcc ctg ggg gcc ggc tca gcc ctg ctg ggg aca gac gtg cag gtg			1056
119	Leu Ala Leu Gly Ala Gly Ser Ala Leu Leu Gly Thr Asp Val Gln Val			
120	121 340 345 350			
122	gaa gcc gca ccc gcc gcc ctg gac tcc gtc ccc tcc tcc gtc gtc cag			1104
123	Ile Ala Ala Ile Ala Ala Leu Val Gln Cys Ile Ser Ser Val Gln			
124	125 355 360 365			

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/655,160

DATE: 01/16/2003

TIME: 13:57:46

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF4\01162003\I655160.raw

127	agt	gac	gag	age	ctc	gac	ctc	age	atc	cag	aac	ggc	ggt	gg:	tca	ggc		1152
128	Ser	Asp	Glu	Ser	Ieu	Asp	Leu	Ser	Ile	Gln	Asn	Arg	Gly	Gly	Ser	Gly		
129	370					375				380								
131	ctg	cag	gcc	gac	tac	agc	atc	ng	gcc	ctg	ggc	gag	gag	ccq	gcc	sga		1200
132	Ieu	Glu	Ala	Ala	Tyr	Ser	Ile	Val	Ala	Leu	Gly	Glu	Glu	Pro	Ala	Arg		
133	385						390			395							400	
135	ggc	gtg	cac	ccg	ctc	tgc	ccc	tgc	gac	acc	gag	atc	ttc	cc:	ggc	aac		1248
136	Ala	Val	His	Pro	Leu	Cys	Pro	Ter	Asp	Thr	Glu	Ile	Phe	Pro	Gly	Asn		
137									405		410						415	
139	ggg	gac	tgc	tac	ccg	ctg	gtg	ttg	gac	aag	ggc	gcc	tgg	ctc	cag	ggg		1296
140	Gly	His	Cy.	Tyr	Arg	Ieu	Val	Val	Glu	Lys	Ala	Ala	Trp	Ieu	Gln	Ala		
141							420			425							430	
143	cag	cag	cac	tgt	tac	ccg	ggc	ttg	ggc	ggc	ggc	atc	ctg	gac				1344
144	Gln	Gly	Gly	Cys	Gln	Aia	Trp	Aia	Gly	Ala	Ala	Leu	Ala	Met	Val	Asp		
145						430			440								445	
147	agt	tct	gc:	gtg	cag	ccg	tcc	ttc	ggg	gtc	acc	agg	agg	ttt				1393
148	Ser	Pro	Ala	Val	Gln	Arg	Leu	Val	Ser	Arg	Val	Thr	Arg	Ser	Ieu			
149						450			455			460						
151	gac	gtg	tgc	atc	ggc	ttt	tcc	act	gtg	cag	ggg	gtg	ggc	gg:				1440
152	Asp	Val	Trp	Ile	Gly	Phe	Ser	Thr	Val	Gln	Gly	Val	Glu	Val	Gly	Pro		
153						465			470		475		480					
155	ggc	ccg	cag	ggc	ttc	tcc	atc	gtg	cag	gg	tgc	cag	aat	tgg	ctg			1488
156	Ala	Pro	Gly	Gly	Gly	Ala	Phe	Ser	Ieu	Glu	Ser	Cys	Gln	Asn	Trp	Ieu		
157						485			490								495	
159	ccg	ggg	ga:	aa:	ccg	ccg	ggc	ccg	gg	ccg	ccg	gg	ccg	ccg	gg:			1536
160	Pro	Gly	Glu	Pro	His	Pro	Ala	Thr	Ala	Glu	His	Cys	Val	Arg	Ieu	Gly		
161						500			505								510	
163	ccc	acc	ggg	tgg	tgt	aa:	ccg	ccg	ttc	ggc	ccg	ccg	ccg	ccg	ttc	cc:		1584
164	Pro	Thr	Gly	Trp	Cys	Asn	Thr	Asp	Ieu	Cys	Ser	Ala	Pro	His	Ser	Tyr		
165						510			520								525	
167	gtc	tgc	ga:	ctg	cag	ccg	ggg	ccg	gtc	ccg	ccg	ccg	ccg	ccg	ccg	cc:		1632
168	Val	Cys	Gly	Ieu	Gln	Pro	Gly	Pro	Val	Gln	Asp	Ala	Glu	Asn	Ieu			
169						530			535								540	
171	ctc	gtc	gg:	gg:	ccc	agt	ggg	ttc	ctg	cag	ggg	ccc	ctg	acc	ctc	ctg		1680
172	Ieu	Val	Gly	Ala	Pro	Ser	Gly	Asp	Ieu	Gln	Gly	Pro	Ieu	Thr	Pro	Ieu		
173						545			550		555						560	
175	gca	cag	ca:	gac	ggg	ttc	ccg	cac	ggg	ccg	ccg	gtg	gg:	gtg	ttg	ttg		1728
176	Ala	Gln	Gly	Asp	Gly	Ieu	Ser	Ala	Pro	His	Glu	Pro	Val	Glu	Val	Met		
177						565			570								575	
179	gtt	tta	cc:	gg:	ttc	ccg	ttg	ttc	ccg	ccg	ttc	cc:	ttt	act	agg	gg:		1776
180	Val	Phe	Pro	Gly	Ieu	Arg	Ieu	Ser	Arg	Glu	Ala	Phe	Ieu	Ter	Thr	Ala		
181						580			585								594	
183	gaa	ttt	gg:	acc	cag	ccg	ggg	ccg	ccg	cag	ccg	ccg	ccg	ccg	ccg	cc:		1824
184	Glu	Phe	Gly	Thr	Gln	Glu	Leu	Arg	Arg	Pro	Ala	Gln	Ieu	Arg	Ieu	Gln		
185						595			600								605	
187	gtg	tac	cc:	cc:	ctc	ccg	acc	ccg	cc:		1872							
188	Val	Tyr	Arg	Ieu	Ieu	Ser	Thr	Ala	Gly	Thr	Pro	Glu	Asn	Gly	Ser	Glu		
189						610			615								620	
191	cc:	ccg	acc	ccg	ttc	ccg	cc:		1920									

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/655,160

DATE: 01/16/2003

TIME: 13:57:46

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF4\01162003\I655160.raw

192 Pro Glu Ser Arg Ser Pro Asp Asn Arg Thr Gin Leu Ala Pro Ala Cys  
 193 625 630 635 640  
 195 atg cca ggg aga cgc tgg tgc cat gga gcc aac atc tgc ttg cca ctg 1968  
 196 Met Pro Gly Gly Arg Trp Cys Pro Gly Ala Asn Ile Cys Leu Pro Leu  
 197 645 650 655  
 198 gac gcc tcc tgc cat ccc tag gtc tgc gcc aat ggc tcc acg tca ggg 2016  
 199 Asp Ala Ser Cys His Pro Gln Ala Cys Ala Asn Gly Cys Thr Ser Gly  
 200 660 665 670  
 201 cca cgg cta ccc ggg gcc acc tat gcg cta tgg aga ggc ttc ctc ttc 2064  
 202 Pro Gly Leu Pro Gly Ala Pro Tyr Ala Leu Trp Arg Glu Phe Leu Phe  
 203 675 680 685  
 204 tcc ttg ctt ccgtt cgg agt ccc acc ggc tcc tcc tcc tgg atc ctc ctc ggc 2112  
 205 Ser Val Pro Ala Gly Pro Pro Ala Glu Tyr Ser Val Thr Leu His Gly  
 206 690 695 700  
 207 cag gag gtg ctc atc ctc ccc ggt gac ctc gtt ggc tgg cag cac gag 2160  
 208 Gln Asp Val Leu Met Leu Pro Gly Asp Leu Val Gly Leu Gln His Asp  
 209 705 710 715 720  
 210 gct ggg cct ggc gca ctc ctc ctc tcc tgg tgg gct ctc ggc cat ctt 2208  
 211 Ala Gly Pro Gly Ala Leu Leu His Cys Ser Pro Ala Pro Gly His Pro  
 212 725 730 735  
 213 ggt ccc cgg gcc ccc tac ctc tcc gcc aac gcc tcc tgg tgg ctc ccc 2256  
 214 Gly Pro Arg Ala Pro Tyr Leu Ser Ala Asn Ala Ser Ser Trp Leu Pro  
 215 740 745 750  
 216 tac ttg cca gct cat ctg gag ggc act tgg ggc tcc tgg atc gcc tgg gcc 2304  
 217 His Leu Pro Ala Glu Leu Gly Thr Trp Gly Cys Pro Ala Cys Ala  
 218 755 760 765  
 219 ctc cgg ctg ctt gca caa cgg gaa cag ctc acc gng ctt ctg ggg ttg 2352  
 220 Ser Leu Arg Leu Leu Ala Gln Arg Glu Glu Leu Thr Val Leu Leu Gly Leu  
 221 770 775 780  
 222 agg ccc aac cat ggc ctg cgg ctg cct ggg cgc tat gag gtc cgg gca 2400  
 223 Arg Pro Asn Pro Gly Leu Arg Leu Pro Gly Arg Tyr Glu Val Arg Ala  
 224 785 790 795 800  
 225 gag gtc ggc aat ggc gtg tcc agg cac aac ctc tcc tcc agc ttt gag 2448  
 226 Glu Val Gly Asn Gly Val Ser Arg His Asn Leu Ser Cys Ser Phe Asp  
 227 805 810 815  
 228 gtg gtc tcc ccc gtg gct ggg ctg cgg gtc aac tac cat gcc ccc cgg 2496  
 229 Val Val Ser Pro Val Ala Gly Leu Arg Val Ile Tyr Pro Ala Pro Arg  
 230 820 825 830  
 231 gac ggc cgc ctc tac gtg ccc acc aac ggc tcc gec tgg tcc tgg ctc gag 2544  
 232 Asp Gly Arg Leu Tyr Val Pro Thr Asn Gly Ser Ala Leu Val Leu Gln  
 233 835 840 845  
 234 gtg gag tct ggc gcc aac gcc acg gec get cgg tgg cct ggg ggc 2592  
 235 Val Asp Ser Gly Ala Asn Ala Thr Ala Thr Ala Arg Trp Pro Gly Gly  
 236 850 855 860  
 237 agt ctc agc gec cgc ttt gac aat gtc tgc cat gec ctc gtc gtc acc 2640  
 238 Ser Leu Ser Ala Arg Phe Glu Asn Val Cys Pro Ala Leu Val Ala Thr  
 239 865 870 875 880  
 240 ttc atg ccc gec tgc cat tgg gac aac gat acc ctg ttc tca gtg 2688  
 241 Phe Val Pro Ala Cys Pro Trp Glu Thr Asn Asp Thr Leu Phe Ser Val

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/655,160

DATE: 01/16/2003

TIME: 13:57:46

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF4\01162003\I655160.raw

157	885	890	895	
158	gtt gca ctg ccc tgg ctc agt gag ggg qaq cac gtg gtg cac gtg gtg			2736
159	Val Ala Leu Pro Trp Ieu Ser Glu Gly Glu His Val Val Asp Val Val			
160	900	905	910	
161	gtg gaa aat agg gtc aag cgg pcc eac ctc aag ctg cgg gtc aeg gca			2744
162	Val Glu Asn Ser Ala Ser Arg Ala Asn Leu Ser Leu Arg Val Thr Ala			
163	915	920	925	
164	gag gag ccc atc ttt ggc ctc ccc pcc aeg ccc aag ccc gag gac agt			2753
165	Glu Glu Pro Ile Cys Gly Leu Arg Ala Thr Pro Ser Pro Glu Ala Arg			
166	930	935	940	
167	gtt gca ctg ccc gtc cta gtg aag tac acc ccc atc gtc aag gac ggc			2762
168	Val Leu Gln Gly Val Leu Val Arg Tyr Ser Pro Val Val Glu Ala Gly			
169	945	950	955	960
170	tcc gag atg ptc ttc cgg tgg aac atc aac gag aag gag tcc atc aac			2771
171	Ser Asp Met Val Phe Arg Trp Ile Asn Asp Lys Gln Ser Leu Thr			
172	965	970	975	
173	tcc cag aac gng gtc ttc aat ptc att tat cag aac pcc pcc gtc ttc			2780
174	Phe Gln Asn Val Val Phe Asn Val Ile Tyr Gln Ser Ala Ala Val Phe			
175	980	985	990	
176	aag ctc tca cng aag gcc tcc aac ccc gtg aac aac gtc acc gtg aac			2789
177	Lys Leu Ser Leu Thr Ala Ser Asn His Val Ser Asn Val Thr Val Asn			
178	995	1000	1005	
179	tcc aac gta aac ttt gag gag ccc aac aac gag aac atc ctc gag			2798
180	Tyr Asn Val Thr Val Glu Arg Met Asn Arg Met Gln Gly Leu Gln			
181	1010	1015	1020	
182	gtc tcc aua ttt gag ggg gtc tcc ccc aat gcc aac ctc gca			2807
183	Val Ser Thr Val Pro Ala Val Leu Ser Pro Asn Ala Thr Leu Ala			
184	1025	1030	1035	
185	ctg aag gcg tgg ctc gtg gtc aac tcc gag pcc pcc gag gtc gac ttc			2816
186	Leu Thr Ala Gly Val Leu Val Asp Ser Ala Val Glu Val Ala Phe			
187	1040	1045	1050	
188	ctg tgg aac ttt tgg cat ggc gag cag gcc ccc ccc aac ttc ccc ccc			2824
189	Leu Trp Thr Phe Gly Asp Gly Glu Gln Ala Leu His Gln Phe Gln			
190	1055	1060	1065	
191	cct ccc tcc aac aac gag tcc ttc ccc gtt ccc gag ccc tcc tgg gtg gcc			2833
192	Pro Pro Tyr Asn Glu Ser Pro Pro Val Pro Asp Pro Ser Val Ala			
193	1070	1075	1080	
194	ctg ctt gtt ttt gtt gag ccc aat gtc aac ccc aac tcc gtt gcc ccc			2842
195	Gln Val Leu Val Glu His Asn Val Thr His Thr Tyr Ala Ala Pro			
196	1085	1090	1095	
197	gtt gag tcc ccc ttt aac atc tcc gca tcc aat gcc ttc gag aac			2851
198	Gly Glu Tyr Leu Leu Thr Val Leu Ala Ser Asn Ala Phe Glu Asn			
199	1100	1105	1110	
200	ctg aac ccc ccc ttt gtg aac gtc gtc gcc tcc ctc ccc tcc			2859
201	Leu Thr Gln Gln Val Pro Val Ser Val Arg Ala Ser Leu Pro Ser			
202	1115	1120	1125	
203	gtg gtc gtg ggt gtg agt gac ggc ctc ctg gtg gtc ggc cgg ccc			2868
204	Val Ala Val Gly Val Ser Asp Gly Val Leu Val Ala Gly Arg Pro			
205	1130	1135	1140	

**RAW SEQUENCE LISTING ERROR SUMMARY**  
PATENT APPLICATION: US/09/655,160

DATE: 01/16/2003  
TIME: 13:57:48

Input Set : A:\PTO.VSK.txt  
Output Set : N:\CRF4\01162003\I655160.raw

**Invalid Line Length:**

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 6

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/655,160

DATE: 01/16/2003

TIME: 13:57:48

Input Set : A:\PTO.VSK.txt

Output Set: N:\CRF4\01162003\I655160.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:34 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:32